

ABSTRACT

The present invention is directed to a systematic *in silico* method to identify new coding sequences, including homologs of coding sequences, in
5 *S. cerevisiae* and other organisms. The present invention is also directed to novel ORFs and the proteins encoded thereby identified using the *in silico* methods.